1200 utility weights associated with cardiovascular events and procedures. Stroke accounted for one-third of total identified cardiovascular utilities, followed by myocardial infarction (25%), heart failure (17%), and peripheral vascular disease (8%). Most (86%) of the utility estimates were derived from secondary references (e.g., published literature). Over one-third (36%) of utilities identified were elicited using EQ-5D and 14% were estimated with direct time trade-off questions. Among the utilities from published studies disclosing sample population information, nearly two-thirds (64%) were elicited from patients and 25 % from community members. Few studies (n=27 CUAs, 168 utility weights) reported utilities for asymptomatic or symptomatic states prior to the cardiovascular event or procedure. **CONCLUSIONS:** Heterogeneity exists in the reporting of cardiovascular utility weights. Analysts conducting CUAs using secondary references can improve study transparency by reporting relevant utility information, such as sample population, sample size, and assessment method. In order to better understand the cost-effectiveness of interventions for cardiovascular conditions, further research is needed to inform baseline utility measurement prior to cardiovascular events or procedures.

PRM117

THE DEVELOPMENT AND VALIDATION OF A REVISED VERSION OF THE MEDICAL OUTCOMES STUDY COGNITIVE FUNCTIONING SCALE (MOS-COG-R)

Yarlas A, White MK, Bjorner JB
Optum, Lincoln, RI, USA

OBJECTIVES: Evaluate the psychometric properties of the revised version of the Medical Outcomes Study Cognitive Functioning Scale (MOS-Cog-R) using data from a representative sample of U.S. adults. **METHODS:** The 6-item MOS-Cog yields a single score representing impairment across a range of cognitive functions, including memory, reasoning, attention/concentration, and confusion, over the previous four weeks. The MOS-Cog-R introduced several changes: one response option was removed; a one-week recall period form was introduced; and 0-100 scoring was replaced by norm-based T-scores (mean=50, standard deviation=10), standardized using data from a 2009 U.S. internet-based general population normative survey to improve interpretability. The psychometric properties of both one-week (acute) and four-week (standard) recall forms of the MOS-Cog-R were examined within the development sample. Scale reliability was evaluated using inter-item correlations and Cronbach's alpha. Construct validity was tested using correlations with relevant continuous criterion variables and analysis of covariance models comparing mean MOS-Cog-R scores across levels of categorical criterion variables. **RESULTS:** Acute and standard forms of the MOS-Cog-R were completed by 2012 and 2003 respondents, respectively. Both forms showed adequate internal consistency reliability (all inter-item correlations>0.40; both Cronbach's alphas≥0.89). MOS-Cog-R scores showed greater correspondence with mental domains of quality of life than physical domains (e.g., correlations for both forms with SF-12v2 mental and physical summary scores were ≥0.55 and ≤0.28, respectively). Direction and magnitude of correlations with criterion measures of psychological status (e.g., depression, anxiety) and health outcomes indicated good convergent validity for both forms. Differences in mean scores across respondents stratified by criterion outcomes supported adequate discriminant validity of each form. **CONCLUSIONS:** The MOS-Cog-R introduces a number of improvements, including simplified responses, the introduction of a one-week recall form, and norm-based scoring that enhances interpretability of scores. Both acute and standard recall-period forms of the MOS-Cog-R demonstrated good reliability and validity.

PRM118

ESTIMATING AN EQ-5D-3L VALUE SET IN SINGAPORE

Luo N, Wang P
National University of Singapore, Singapore, Singapore

OBJECTIVES: To establish an EQ-5D-3L value set using time trade-off (TTO) values directly measured from the general Singaporean population. **METHODS:** The values of 80 EQ-5D-3L health states were directly elicited from a general Singaporean population sample using a TTO method modified from the MVH protocol. In face-to-face interviews, participants were asked to value a block of 10 health states. Various linear regression models and model specifications were examined to assess their goodness of fit to the data, at both individual and aggregated levels, and ability to predict the values for measured and unmeasured EQ-5D-3L health states. Goodness of fit was assessed in terms of mean absolute error (MAE), numbers of prediction errors larger than 0.05 and 0.10; while prediction ability was assessed in terms of logic consistency and bias. **RESULTS:** A total of 500 participants provided data for this study. The N3 model without a constant using the random-effects estimator exhibited the best fit of the data at individual level, predicted values with the least bias, and generated logically consistent values for all 243 EQ-5D-5L health states. The MAE was 0.1227, and 41 out of 80 predicted values had errors less than 0.10 in absolute magnitude. Based on this model, the second highest utility value is 0.8867 for state 21,111and the lowest value is -0.7284 for state 33,333. **CONCLUSIONS:** This study established the value set of EQ-5D-3L health states for Singapore. The value set provides health services researchers in Singapore a useful tool for assessing the cost-effectiveness of health technologies and services.

PRM119

VALIDATION AND PSYCHOMETRIC EVALUATION OF A HEALTH CARE ORIENTATION ASSESSMENT

Bushnell DM[1], McCarrier KP[1], Scanlon M[1], Nelson DR[2], Martin ML[1], Buesching DP[2]
[1]*Health Research Associates, Inc., Seattle, WA, USA,* [2]*Eli Lilly and Company, Inc., Indianapolis, IN, USA*

OBJECTIVES: The Provider-Dependent Health Care Orientation (PDHCO; Kaplan 1996) assesses an individual's orientation towards health and health care and measures an individual's dependence (i.e., passivity) related to health care and disease management. We sought to build on prior validation of the instrument by evaluating the reproducibility of the PDHCO and testing equivalence between paper and electronic administration modes. **METHODS:** The PDHCO and other questionnaires were administered to a sample of adults recruited through web-based advertisements in 8 U.S. cities. Participants completed the PDHCO on paper and computerized formats in a randomized crossover design. A one-week retest was completed at home. Reproducibility and mode equivalence were assessed using the intraclass correlation coefficient (ICC). Cronbach's alpha was calculated to assess internal consistency. To assess convergent validity, the correlation of the PDHCO to the Communication With Physician (CWP) Scale of the Chronic Disease Self-Efficacy Scale was calculated. ANOVA was used to compare mean PDHCO scores among tertiles of the Health Assertiveness Scale (HAS). **RESULTS:** Of the 230 participants that completed baseline assessment, 228 (99.1%) completed the one-week retest. The mean age of participants was 44.3 years, 51.3% were female, and 58.3% were Caucasian. A small number (n=9; 3.9%) reported their health as Poor. The mean PDHCO score was 49.7(±14.7), and the ICC between paper and computerized administration was 0.887. The ICC for the one-week retest of the paper format was 0.913, and the PDHCO was found to be internally consistent (Cronbach's alpha=0.735). Significant correlations were found with the CWP (r=-0.246; p<.001), and the instrument discriminated between levels of the HAS (p<0.05). **CONCLUSIONS:** The PDHCO was observed to have adequate reproducibility and internal consistency as well as appropriate convergent validity. The scale was found to significantly discriminate between levels of health assertiveness. Equivalence between paper and web-based administration was demonstrated.

PRM120

DEVELOPMENT OF A NEW PATIENT-REPORTED OUTCOME (PRO) INSTRUMENT FOR PULMONARY ARTERIAL HYPERTENSION (PAH): THE PULMONARY ARTERIAL HYPERTENSION–SYMPTOMS AND IMPACT (PAH-SYMPACT) QUESTIONNAIRE

McCollister D[1], Kummer S[2], Badesch DB[1], Filusch A[3], Hunsche E[4], Schüler R[5], Wiklund I[6], Peacock A[7]
[1]*University of Colorado Denver, Aurora, CO, USA,* [2]*United BioSource Corporation, Bethesda, MD, USA,* [3]*HPK - Heidelberg Private Clinic, Heidelberg, Germany,* [4]*Actelion Pharmaceuticals Ltd., Allschwil, Switzerland,* [5]*Actelion Pharmaceuticals UK Ltd., London, UK,* [6]*United BioSource Corporation, London, UK,* [7]*Pulmonary Vascular Unit, Regional Heart and Lung Centre, Glasgow, UK*

OBJECTIVES: In the absence of any pulmonary arterial hypertension (PAH)-specific PRO instruments developed in accordance with 2009 FDA guidance requirements, this qualitative research study was conducted to develop a new instrument assessing PAH symptoms and their impacts following the PRO guidance. **METHODS:** Patient inclusion criteria were age 18–80 years and symptomatic PAH (WHO Group 1) diagnosed by right-heart catheterization. Concept elicitation was based on 5 focus groups, after which saturation of emergent concepts was reached. A PRO instrument for PAH symptoms and their impacts was drafted, considering input from the international Steering Committee as well as translatability and legibility assessments. Two rounds of cognitive interviews on the draft PRO were conducted, with instrument revisions following each. The study was approved by institutional review boards at 5 US sites and participants provided written informed consent. **RESULTS:** Focus groups comprised 25 patients, and 20 additional patients participated in cognitive interviews (10 per round). Participants had a mean±SD age of 54±16 years, were predominantly female (91%), and were diverse in race/ethnicity, WHO functional class (FC I/II: 51%, III/IV: 49%), and etiology (associated PAH: 51%, idiopathic PAH: 47%, familial PAH: 2%). The draft PRO instrument was found to be clear, comprehensive, and relevant to PAH patients in cognitive interviews. Items were organized in a draft conceptual framework with 4 symptom domains (respiratory symptoms, tiredness, cardiovascular symptoms, other symptoms) and 5 impact domains (physical activities, daily activities, social impact, cognition, emotional impact). The recall period is the past 24 hours for symptom items, and 7 days for impact items. **CONCLUSIONS:** The draft instrument was shown to capture symptoms and their impacts relevant to PAH patients, demonstrating content saturation and concept validity. Additional testing is needed to confirm the content and psychometric validity of the PAH-SYMPACT before use in future clinical practice or studies.

PRM121

USING A LIFE SATISFACTION MEASURE IN THE VALUATION OF HEALTH: A CASE STUDY OF MULTIPLE SCLEROSIS

Suponcic S, DiBonaventura M, Kopenhafer L
Kantar Health, New York, NY, USA

OBJECTIVES: The valuation of health is becoming increasingly important for the purposes of health technology assessment. However, recent research has suggested that traditional preference-based measures for valuing health states may be flawed and life satisfaction measures may better represent the patient perspective. The current study extended this research by examining the validity of a life satisfaction measure among patients with multiple sclerosis (MS). **METHODS:** Data from the Multiple Sclerosis Patient study (N=1000) were used, which were collected from an Internet-based survey of patients who self-reported a diagnosis of MS. Information on demographics, disease and treatment history, and health outcomes were collected. A life satisfaction measure was included (1=not satisfied to 7=completely satisfied) and examined alongside MS-specific symptom and quality of life measures. **RESULTS:** Respondents were mostly female (82.8%) with an average age of 48.7 years. Most respondents had relapsing remitting MS (89.0%). The mean life satisfaction scores were 4.87 (SD=1.26; IQR:4-6). Only small effects were observed between life satisfaction and

symptom-related disability measures such as the Patient-Determined Disease Steps measure (r=-0.19) and the Multiple Sclerosis Rating Scale-Revised (r=-0.26). Although similarly-sized effects were observed with activities of daily living and symptom subscales of the MusiQoL (rs = 0.31, both p<0.05), much stronger relationships with psychological well-being (r=0.47), sentimental and sexual life (r=0.44), and coping (r=0.42; all p<0.05) were observed. **CONCLUSIONS:** Disability, symptom, and activity measures were only modestly related to life satisfaction among patients with MS. Conversely, psychological, emotional, and sexual factors were much more strongly associated. These results suggest that the domains of traditional preference-based measures may not ultimately focus on those that matter most for the patient. Further research should continue to examine the importance of alternative measures to most accurately value health states.

PRM122

RESULTS FROM A D-EFFICIENT DISCRETE CHOICE EXPERIMENT DESIGN AND PROTOCOL WITH CHOICE SETS OF 3 STATES EACH TO ELICIT AN EQ-5D-3L VALUE SET

Bailey H

The University of the West Indies, St. Augustine, Trinidad and Tobago

OBJECTIVES: Discrete Choice Experiments (DCEs) have received considerable interest in EQ-5D valuations because these methods are simple and choice based. Some investigations using DCEs to develop EQ-5D Value Sets include implausible states. The DCE literature discourages the use of implausible options, citing the links between cognitive burden (facing respondents) and random variability in responses. It has been suggested that giving DCE respondents choice sets with more than 2 alternatives may improve respondent effort/concentration. **METHODS:** A D-Efficient DCE design for EQ-5D-3L comprising 15 choice sets of 3 alternatives each, that excluded states which combine a '3' on Mobility with a '1' on either Usual Activities, or Self Care or both was developed. As a pilot test, and as a first step towards developing an on-line elicitation exercise, an elicitation exercise was put into a Microsoft PowerPoint file which was sent out to 189 respondents at various work locations in Trinidad and Tobago with a submission deadline of 2 weeks. The elicitation process was based on a protocol using sequential comparisons that break the choice process into s series of simple ranking steps. **RESULTS:** Sixty-five Respondents returned DCE results within the 2 week period of which the results from 43 respondents were useable (66%). A utility function was estimated based on data from the 43 useable files. This produced an internally valid model with all coefficients bearing the correct signs, and relative magnitudes, and all significant to the 5% level. **CONCLUSIONS:** This DCE design allows for EQ-5D-3L valuation studies that can be conducted with small samples (no blocking) and it avoids the use of implausible states. It introduces a DCE protocol that allows respondents to easily make choices out of choice sets comprising 3 EQ-5D-3L States. Increasing the number of alternatives in each choice set allows for designs with fewer choice sets.

PRM123

HEALTH RELATED QUALITY OF LIFE (HRQOL) OF ARGENTINE POPULATION: RESULTS FROM THE NATIONAL HEALTH RISKS FACTOR SURVEY

Augustovski F[1], Rey-Ares L[2], Gibbons L[3]

[1]*Health Economic Evaluations and Technology Assessment at the Institute for Clinical Effectiveness and Health Policy, Buenos Aires, Argentina,* [2]*Institute for Clinical Effectiveness and Health Policy (IECS), CABA, Argentina,* [3]*Institute for Clinical Effectiveness and Health Policy, Buenos Aires, Argentina*

OBJECTIVES: To describe general population self reported visual analog scale (SR-VAS), as well as time-trade off (TTO) and visual analog scale (VAS) preference values (PV) by age groups and gender. **METHODS:** In 2009, the second "Health Risks Factors National Survey" was undertaken in Argentina, including 34,728 subjects, age 18-65 yrs, randomly selected from all Argentine provinces with a probabilistic multi-stage sample design. The expanded population was 24,434,595. Data were obtained on sociodemographics, health risks factors, and health status (EQ-5D and the general health question of SF-36). PV were assigned using weights derived from a previous local study. **RESULTS:** Over 80% of the population reported being healthy (11.19% reported health status as excellent, 26.88% as very good and 42.68% as good), whereas 19.25% of the population reported a regular or poor health. The presence of limitations in each EQ-5D domains varied from 2.28% in personal care to 30.15% in Pain/Discomfort. The presence of limitations was higher in women (3.14% in personal care and 33.38 in Pain/Discomfort). The population-weighted mean of the SR-VAS was 76.49 (95%CI 76.17-76.82). Utilities of the EQ-5D based on preferences in Argentina were 0.881 (95%CI 0.878-0,884) for VAS and 0.91 (95%CI% 0.907-0.913) for TTO. All values were systematically lower in women. Regarding the differences between age groups, values tended to decrease consistently with the increasing of age. There were no significant differences between results of the first survey conducted in 2005 and the second one in 2009. **CONCLUSIONS:** There are few surveys in Latin America that incorporated the EQ-5D tool to describe general population health status and Argentina is one of the few Latin American countries that derived local preferences based on TTO and VAS methods. These results can serve as a benchmark for future population studies and also as inputs for cost-utility analysis of health technologies.

PRM124

EXAMINATION OF PREFERENCES ELICITED FROM THE GENERAL PUBLIC: SUBGROUP COMPARISONS OF RESULTS FROM THE TIME TRADE-OFF AND STANDARD GAMBLE INSTRUMENTS

Iskedjian M[1], Navarro V[2], Farah B[3], Berbari J[3], Walker JH[4], Le Lorier J[1]

[1]*Université de Montréal, Montréal, QC, Canada,* [2]*PharmIdeas Europe SAS, Lyon, France,* [3]*PharmIdeas Research and Consulting, Ottawa, ON, Canada,* [4]*Brock University, St Catharines, ON, Canada*

OBJECTIVES: The time trade-off (TTO) and the standard gamble (SG) are two methods used for the direct measurement of health-related quality of life, expressed as health utilities (HU), among the general public. Significant overall correlation was observed in a previous systematic review comparing TTO and SG values reported by the general public (19 pertinent publications, 77 study arms and 24 diseases). In the present analysis, correlation was further explored among specific demographic, clinical and analytic subgroups. **METHODS:** When a sufficient sample of study arms was available, Spearman's rank correlation test was applied to test the correlation between mean or median TTO and SG by sorting the TTO values in increasing increments of 0.1, based on the following subgroups: incremental HU, disease, gender, age, income, education level and employment status. Significance was defined as p-value<0.05. **RESULTS:** Significant positive correlation was observed between mean TTO and SG values for all increments of TTO except when TTO<0.5 (r = 0.429), for ocular disease (r = 0.976), skin disease (r = 0.964) and arthritis (r = 0.956). A post-hoc analysis quadrupling the sample of studies with TTO<0.5 resulted in significant positive correlation, indicating that the lack of significant correlation was possibly due to the scarcity of studies. The series of subgroup analyses focusing on median incremental HU values or median HU values associated with skin disease resulted in similar findings to the analysis of means. No analysis of demographic subgroups could be undertaken due to the limited study sample. **CONCLUSIONS:** A correlation between TTO and SG outcomes reported by members of the general public was observed for numerous subgroups. It is possible that correlation would have been observed in additional subgroups, but was not, due to the limited study sample.

PRM125

CROSS-CULTURAL ADAPTATION OF THE HAEMO-QOL QUESTIONNAIRE INTO 28 LANGUAGES

Von Mackensen S[1], Gil Campos I[2], Acquadro C[3], Strandberg-Larsen M[4], Fuller DS[5]

[1]*University Medical Centre Hamburg-Eppendorf, Hamburg, Germany,* [2]*MAPI Institute, Lyon, France,* [3]*MAPI Research Trust, Lyon, France,* [4]*Novo Nordisk A/S, Søborg, Denmark,* [5]*MAPI Institute, Philadelphia, PA, USA*

OBJECTIVES: The Haemo-QoL questionnaire was developed to assess health-related quality-of-life (HRQoL) in children and adolescents with haemophilia. Three different age-group versions (4–7, 8–12, 13–16 years) are available as well as three corresponding proxy versions for their parents. The original questionnaire was developed in German. Since capturing HRQoL aspects of haemophilia care has become an integrated part of clinical trials in this field, it is important to ensure that the questionnaires used are linguistically validated for international use. The objective of this study is to present the cross-cultural adaptation of the Haemo-QoL into 28 languages representing six different language families (Indo-European, Ural-Altaic, Afro-Asiatic, Japonic, Sino-Tibetan, and Austronesian). **METHODS:** A standard multi-step methodology was used for translation and cultural adaptation: 1. Development of a concept list; 2. Forward/backward translation (or adaptation step or quality check); 3. Review of the backward translation and report by the developer of the original instrument; and 4. Review of the translation by a clinician. Difficulties encountered during the process were categorized as Grammatical, Idiomatic, Semantic/Conceptual, and Cultural. **RESULTS:** Fifteen items raised discussions for semantic (12 items), cultural (2 items), and idiomatic reasons (1 item). For instance, the statement "I had to refrain from sports like rollerblading or soccer" raised cultural issues. The examples of sports had to be culturally adapted, and became "rugby or soccer" (in South African English), or "ice-skating or sports like soccer" (in Mandarin). Further examples will be presented. **CONCLUSIONS:** The cross-cultural adaptation of the Haemo-QoL into 28 languages required an international collaboration and enabled the production of conceptually equivalent and culturally appropriate tools. The same process was used for the Haem-A-QoL (for adults) and the treatment satisfaction questionnaire (Hemo-Sat). When applied, these linguistically validated tools will provide insights into an area of haemophilia not well understood in the past.

PRM126

BENCHMARKS FOR INTERPRETATION OF SCORE DIFFERENCES ON THE SF-36 HEALTH SURVEY

Bjorner JB[1], Wolden M[2], Gundgaard J[2], Bayliss MS[1], Miller K[1]

[1]*Optum, Lincoln, RI, USA,* [2]*Novo Nordisk A/S, Søborg, Denmark*

OBJECTIVES: Patient-reported outcomes (PROs) are widely used in clinical research. However, results on score differences may be hard to interpret if clinicians are unfamiliar with the assessment tools and lack benchmarks for interpretation of results. This study aims to estimate clinical and social benchmarks for interpretation of score differences on the SF-36 Health Survey and to test whether the interpretation depends on score level and patient background characteristics. **METHODS:** Using survival and logistic regression models, we reanalyzed data from three US cohort studies: the Medical Outcomes Study (N=3,445), the Medicare Health Outcomes Survey (N=78,183), and the QualityMetric 2009 Norming Study (N=4,040). Outcome variables were mortality, hospitalization, current inability to work, and loss of ability to work. **RESULTS:** Benchmarks were robust across types of disease and for people with or without chronic disease. However, the benchmarks did vary according to age and score level. A one-point lower score on the PF, GH and PCS scales was associated with a 1.05-1.09 relative risk (RR) of mortality for the typical chronic disease patient, with stronger associations in the younger age groups. For several scales (PF, RP, BP, GH, VT, SF, and RE), the associations with mortality also depended on score level, with stronger associations in the lower score ranges (i.e. patients in worse health). A one-point lower score on the PF, RP, BP, GH, VT, SF, and